

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-242006

(43)Date of publication of application : 21.09.1993

(51)Int.Cl.

G06F 13/14
G06F 3/06

(21)Application number : 04-078523

(71)Applicant : MITSUBISHI ELECTRIC CORP

(22)Date of filing : 28.02.1992

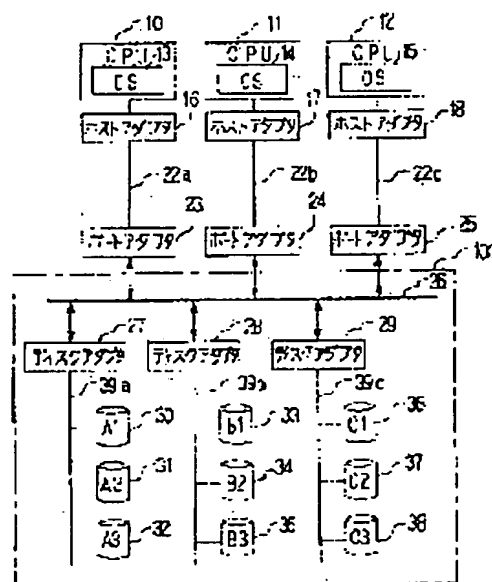
(72)Inventor : NOGUCHI MASAHIRO

(54) MEMORY CONTROL SYSTEM

(57)Abstract:

PURPOSE: To efficiently utilize resources of an external storage device, to discontinue its explicit reserve command and release command, and to preferentially provide resources to a specific CPU.

CONSTITUTION: For instance, a data transfer between a CPU 10 and an external storage device 100 is executed by a packet of length of an integer multiple of a storage unit of the external storage device 100, and as for the packet, there are two kinds of packets for showing a request system and a response system, and as for a request to a designated storage device 30 of the external storage device 100 from the CPU 10, the next request can be issued from the CPU 10 by receiving a response from the designated storage device 30.



LEGAL STATUS

[Date of request for examination]	21.03.1996
[Date of sending the examiner's decision of rejection]	19.05.1998
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	2834927
[Date of registration]	02.10.1998
[Number of appeal against examiner's decision of rejection]	10-09557
[Date of requesting appeal against examiner's decision of rejection]	18.06.1998
[Date of extinction of right]	

Copyright (C); 1998,2003 Japan Patent Office

THIS PAGE BEHIND (verio)

(19) 日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11) 特許出願公開番号
特開平5-242006
(43) 公開日 平成 5 年 (1993) 9 月 21 日

(51) Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 13/14	3 1 0 Y	7230-5B		
3/06	3 0 1 F	7165-5B		

審査請求 未請求 請求項の数 1 (全 11 頁)

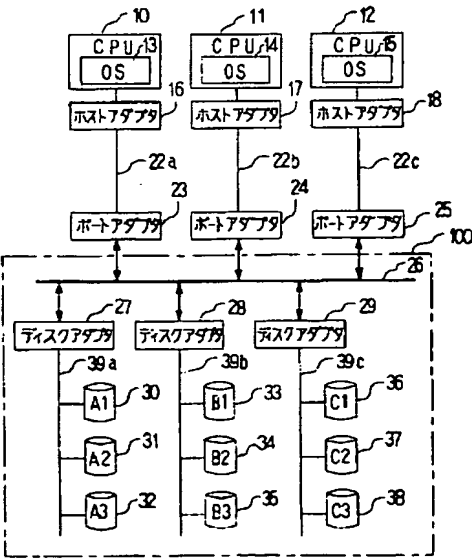
(21) 出願番号	特願平4-78523	(71) 出願人	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目2番3号
(22) 出願日	平成4年(1992)2月28日	(72) 発明者	野口 昌弘 鎌倉市大船五丁目1番1号 三菱電機株式 会社情報電子研究所内
		(74) 代理人	弁理士 宮園 純一

(54) 【発明の名称】 メモリ制御方式

(57) 【要約】

【目的】 外部記憶装置資源の効率的な利用を可能とし、明示的なリザーブコマンドやリリースコマンドの廃止と特定CPUへの優先的なリソースの提供を可能にする。

【構成】 例えばCPU10と外部記憶装置100との間のデータ転送を、外部記憶装置100の記憶単位の整数倍の長さのバケットで行い、該バケットとしてはリクエスト系とレスポンス系を示す2種類のバケットがあり、CPU10から外部記憶装置100の指定記憶デバイス30へのリクエストに対しては該指定記憶デバイス30からのレスポンスの受信により、次のリクエストがCPU10から発行可能となる。



22a~22c: 信号線、26: データバス、
30~38: 外部記憶デバイス、39a~39c: データバス、
100: 外部記憶装置

(2)

特開平 5-242006

1

【特許請求の範囲】

【請求項 1】 複数の計算機の中央処理装置間で外部記憶装置を共有する計算機システムにおいて、当該中央処理装置と上記外部記憶装置との間のデータ転送を、該記憶装置の記憶単位の整数倍の長さのバケットで行い、該バケットとしてはリクエスト系とレスポンス系を示す 2 種類のバケットがあり、当該中央処理装置から上記外部記憶装置の指定記憶デバイスへのリクエストに対しては該指定記憶デバイスからのレスポンスの受信により、次のリクエストが当該中央処理装置から発行可能となるように制御することを特徴とするメモリ制御方式。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は、計算機システムの資源管理方式に係わり、特に複数の計算機間で磁気ディスク等の外部記憶装置を共有する場合に好適な制御を行うメモリ制御方式に関するものである。

【0002】

【従来の技術】 図 6 は例えば公開特許公報平 2-10424 号に示された従来のメモリ制御方式を用いた計算機システムの構成を示すブロック図である。図 6 において 10, 11, 12 は磁気ディスク装置等の補助記憶装置 21a~21c を共有する計算機の CPU、13, 14, 15 は独立の OS (オペレーティングシステム) である。各 CPU 10, 11, 12 は、それぞれチャンネル 116, 117, 118 と接続され、各チャンネル 116, 117, 118 は制御用メモリ 20 を有する補助記憶制御装置 19 を介して被制御対象の補助記憶装置 21a~21c と接続されている。

【0003】 次に動作について説明する。CPU 10, 11, 12 間で共用する補助記憶装置 21a~21c 上のデータに対するアクセスは、1 時点では 1 つの CPU からしかできない。このためユーザは、当該補助記憶装置の使用権を得るための排他制御要求を当該 CPU の OS に対し発行しなければならない。排他制御要求を受けとった OS は、排他制御要求の後の最初のデータ転送入出力要求に対応する CCW (チャンネルコマンド語) の先頭に排他制御用のリザーブコマンドをチェインする。CCW はチャンネルを通り補助記憶制御装置 19 に伝えられる。CCW は例えば図 7 で示す形式になっている。CCW 71 の最初はコマンドコードであり、このコマンドコードは例えばリザーブコマンドが X' 24'、リリースコマンドが X' 44'、無条件リザーブコマンドが X' 54' で表現される。リザーブコマンドは、排他制御識別子 72 をキーとして、補助記憶装置 21a~21c のロックを確保し、他の CPU からの使用を禁止する。リリースコマンドは、排他制御識別子 72 をキーとしたロックを解除する。無条件リザーブコマンドは、排他制御識別子 72 をキーとした他の CPU のロックを強制的に解除し、自 CPU に、排他制御識別子 72 をキーとした

2

ロックを確保するコマンドである。

【0004】 例えばリザーブコマンドを受けとった補助記憶制御装置 19 は、リザーブコマンドに指定された排他制御識別子 72 をキーとし、補助記憶制御装置 19 内の制御用メモリ 20 に記憶された排他制御識別子管理テーブル (図示せず) を参照する。補助記憶制御装置 19 は、排他制御識別子管理テーブルに該排他制御識別子 72 がすでに登録されている場合、その排他制御識別子 72 の登録を行ったチャンネルルートを読みだし、今回発行されたリザーブコマンドのチャンネルルートと一致するか、または同一のチャンネルバスグループに属するかを判定を行う。これにより、他の CPU から当該データセットが他の CPU からロックされていないと判断された場合、排他制御識別子管理テーブルに該排他制御識別子 72 を登録する。すでに他の CPU にロックが確保されている場合は、要求 CPU に対しデバイスの使用中を報告する。

【0005】 また、リリースコマンドは、排他制御識別子管理テーブルから該排他制御識別子 72 の削除を行う。無条件リザーブコマンドは、排他制御識別子管理テーブルに該排他制御識別子 72 の登録を強制的に行う。

【0006】

【発明が解決しようとする課題】 従来のメモリ制御方式は以上のように構成されているので次のような問題点があった。第 1 に、補助記憶制御装置にすべての信号/データが集結しているので性能のボトルネックがここで発生する。第 2 に、補助記憶制御装置により CPU のアクセスがシリアル化されてしまうので、ある CPU が大量のデータ転送を行うとその後に続く CPU のアクセスが待たされることになり 1/0 タイムアウトエラーを誘発することになる。第 3 に、リザーブコマンドを発行した CPU がシステムダウンするとリザーブされた補助記憶装置が使用不能になる。この状態を回復するための無条件リザーブコマンドの発行は、補助記憶装置以外の手段で各 CPU が通信しシステムダウンを起こした CPU を検知/判断しそれを管理する上位 CPU が行う必要があるためのこのコマンド自体の発行が難しい。

【0007】 この発明は上記のような問題点を解決するためになされたもので、外部記憶装置資源の効率的な利用を可能とし、明示的なリザーブコマンドやリリースコマンドの廃止と特定の中央処理装置への優先的な外部記憶装置資源の提供を可能とするメモリ制御方式を得ることを目的とする。

【0008】

【課題を解決するための手段】 この発明に係るメモリ制御方式は、当該中央処理装置と外部記憶装置 100 との間のデータ転送を、該記憶装置 100 の記憶単位の整数倍の長さのバケットで行い、該バケットとしてはリクエスト系とレスポンス系を示す 2 種類のバケットがあり、当該中央処理装置から外部記憶装置 100 の指定記憶デ

(3)

特開平5-242006

3

バスへのリクエストに対しては該指定記憶デバイスからのレスポンスの受信により、次のリクエストが当該中央処理装置から発行可能となるように制御するものである。

【0009】

【作用】例えば中央処理装置11と外部記憶装置100との間のデータ転送は、この記憶装置100の記憶単位の整数倍の長さのバケットで行われる。中央処理装置11が外部記憶装置100の記憶デバイス33へリクエスト系のバケットを発行すると、記憶デバイス33からのレスポンス系のバケットを受信するまでは次のリクエスト系のバケットを発行できない。したがって、中央処理装置11は外部記憶装置100の記憶デバイス33からレスポンス系のバケットを受信すれば、次のリクエスト系のバケットを発行できる。

【0010】

【実施例】図1はこの発明の一実施例に係るメモリ制御方式を用いた計算機システムの構成を示すブロック図である。図1において、図6に示す構成要素に対応するものには同一の符号を付しその説明を省略する。図1で206, 17, 18はCPU10, 11, 12と外部記憶装置100を接続するためのホストアダプタで、CPU10, 11, 12からのコマンドを解析し外部記憶装置100に適した形に処理したり外部記憶装置100からのデータをCPU10, 11, 12に適した形に処理する機能がある。22a, 22b, 22cはホストアダプタ16, 17, 18と外部記憶装置100を接続する信号線で、長距離でのデータ送受信を行うために光ファイバーなどを使用する。23, 24, 25はホストアダプタ16, 17, 18と外部記憶装置100をインターフェイスするポートアダプタで、ホストアダプタ16, 17, 18とは対で使用される。26は外部記憶装置100内のデータバスで、データ転送はバーストで行われ、このデータバス26のデータ転送能力はこの外部記憶装置100に接続されるCPU10, 11, 12のデータ転送率の総和を越える高速タイプである。27, 28, 29は磁気ディスク等の外部記憶デバイスを制御するための外部記憶制御装置であるディスクアダプタで、SCSI等の外部記憶デバイス制御のためのデータバス39を生成/制御する。30から38は外部記憶デバイスで、A1からC3までそれぞれユニークにアドレッシングされている。ホストアダプタ16, 17, 18、ポートアダプタ23, 24, 25、ディスクアダプタ27, 28, 29も同様にユニークにアドレッシングされている。

【0011】図2は図1中のホストアダプタの構成を示すブロック図である。図2において、200はシステムバス204や内部バス205とのデータの入出力を行うシステムバスインターフェイス、201はマイクロプロセッサ、202はDMA（直接メモリアクセス）コント

4

ローラ、203はポートアダプタとのデータの入出力を行うフロントインターフェイスである。システムバス204は計算機のCPU10, 11, 12（図1参照）に接続されている。

【0012】図3は図1の実施例においてホストアダプタからディスクアダプタ間で使用しているバケットの形式を示したもので、バケット形式は(a), (b), (c)のように3つの形式がある。まず(a)のHeadingは目的のエレメントへ制御/状態情報を伝達するために使用するバケットである。ここで40は目的エレメントのアドレスを格納するためのエリアで、各エレメントはこの値を参照し自分のアドレスと一致した場合、このバケットをとり込む。自アドレスに一致しないバケットはポートアダプタの場合、ホストアダプタから送出されたデータは内部データバス26へ送出し、内部データバス26上のデータは無視する。ディスクアダプタの場合、自アドレスに一致しないバケットは無視する。41はソースアドレスでバケット送出元のアドレスが格納されている。レスポンス系のバケットはリクエスト系バケットのこのアドレスを目的（ディストネーション）アドレスとして再使用する。42はディスク情報でディスクの制御情報やステータスが格納される。43はバケットタイプで、ここでこのバケットが何であるか例えばリクエスト系/レスポンス系、リード/ライトなどが示される。

【0013】(b)のHeading+1Logical Block DataはHeading情報/ステータスとともに外部記憶装置100からの/外部記憶装置100へのデータをディストネーションへ送るためのもので外部記憶装置100の1単位長のデータ44を送る。CPUからのデータ書き込みの場合、1単位長に満たないデータについてその不足分を埋めるのはホストアダプタの責任である。

【0014】(c)のHeading+2Logical Block DataはHeading情報/ステータスとともに外部記憶装置100からの/外部記憶装置100へのデータをディストネーションへ送るためのもので外部記憶装置100の2単位長のデータ45を送る。なお、図3における40, 41, 42, 43の内容は図2のマイクロプロセッサ201により生成される。また、図3における44, 45の内容は図2のDMAコントローラ202により計算機（CPU）から直接データが受信される。

【0015】この実施例におけるデータの転送はこのHeading+1Logical Block DataとHeading+2Logical Block Dataの組合わせで実行される。

【0016】この実施例のメモリ制御方式の特徴は、当該CPUと外部記憶装置100との間のデータ転送を、該記憶装置100の記憶単位の整数倍の長さのバケット

(4)

特開平5-242006

5

で行い、該パケットとしてはリクエスト系とレスポンス系を示す2種類のパケットがあり、当該CPUから外部記憶装置100の指定記憶デバイスへのリクエストに対しては該指定記憶デバイスからのレスポンスの受信により、次のリクエストがCPUから発行可能となるように制御する。

【0017】なお、上記記憶単位とはアクセス対象の記憶装置に対する1回のアクセスで読み書きできる単位あるいはアドレス等により特定できる記憶単位を言う。例えば計算機の主記憶装置であれば8ビットを単位としたバイトが記憶単位となる。外部記憶装置は一般に大容量であるためデータの記憶(WRITE)/読みだし(READ)は計算機の主記憶装置よりも大きい単位で行う。この単位がディスクドライブではセクターと呼ばれ、単位は例えば商品名MELCOM70/MXのコンピュータでは1024バイトである。よって本実施例では1024バイトを単位としてパケットを使用する。

【0018】次に図2に示すホストアダプタの動作を説明する。例えばCPUからのデータを外部記憶装置に転送する場合、マイクロプロセッサ201は図3に示すようなパケットのヘディング部(Heading)を生成し内部バス205を介してフロントインターフェイス203へ渡す。次にマイクロプロセッサ201はシステムバス204からシステムバスインターフェイス200を経由して内部バス205にデータを流すようにシステムバスインターフェイス200に指示する。同時にマイクロプロセッサ201は、DMAコントローラ202に対し上記データをマイクロプロセッサ201を経由せずに直接フロントインターフェイス203にセットするように指示する。これによりフロントインターフェイス203はパケットのデータ部を高速に次段のポートアダプタへ渡す。このようにしてデータパケットが生成される。

【0019】外部記憶装置のデータをCPUに転送する場合、フロントインターフェイス203から入力データが内部バス205に流されると、マイクロプロセッサ201がヘディング部からパケットの種類とデータの長さを解析しDMAコントローラ202を使ってデータ部をシステムバスインターフェイス200およびシステムバス204を経由してCPUに伝える。

【0020】図4はこの実施例においてパケットがどのように実行されるかを示したプロトコルフロー図であり、CPU10、ホストアダプタ16、ポートアダプタ23を経由して並列動作でディスクB1(Disk B1)とディスクC1(Disk C1)にリード起動をかけた場合を示す。図5はノーオペレーション(No op)パケットを使ってデータバス系試験を実行している様子を示すプロトコルフロー図である。なお、図4および図5において、ディスクB1、C1は図1における外部記憶デバイス33、36に相当する。

【0021】次に図1の実施例の動作について図4によ

6

り説明する。CPU10からリード起動を外部記憶デバイスB1へかけるとホストアダプタ16はCPU10内の制御情報(例えばCCW)を読みだし外部記憶装置100に適したパケット形式へ変換する。このパケットはポートアダプタ23へ信号線22aを経由して送出される。Read Request Packet(a1, B1)がそれである。このパケットは上述したように最大長が規定されているのでデータバスを長時間占有することがないため、CPU10およびホストアダプタ16はただちにディスクC1へのリード起動(Read Request Packet(a2, C1))をディスクB1と同様に実行できる。これらのパケットはポートアダプタ23に対するものではないのでそのまま内部バス26へ送出される。ディスクアダプタ28、29はこれらパケットが自分の管理する外部記憶デバイスに対するものである場合、該当するパケットを取り込み制御情報を解析して該当ディスクへ送る。

【0022】この図4ではディスクB1よりディスクC1の方が先に動作を終了しているため、ディスクアダプタ28はリクエストパケットの動作終了を示すレスポンスパケットを送出している。Read Response Packet(a2, C1)がそれである。このように各パケットはREAD/RESPONSEが対で実行されることとパケットの中にソースとディストネーションの各アドレスが格納されているため、このようにパケットの順序の入れ替わりが可能となっている。ディスクB1へのアクセスは1回のパケット転送では終了しなかったため、ひきつづきRead Response Packet(a1, B1)受信後Read Request Packet(a1, B1)の送信が行われた。このようにディスクアダプタはレスポンスを返すタイミングを制御できるため公平な資源のアクセス、あるいは優先的資源のアクセスが可能となる。

【0023】この図4においてポートアダプタ23とディスクアダプタ28の間のパケット列は内部バス26のトラフィックを示すものでありここでも特定データのバスの占有がないことがわかる。また、ここで外部記憶装置100内で使用されているパケットはCPUのコマンド形式に左右されない外部記憶に適した形式となっているため外部記憶装置100として資源の利用効率が高い。

【0024】次に図5によりオンライン中のデータバス系チェックの方法を説明する。この図5はチェックシーケンスのみをぬきだして書いているが、このチェックに使うコマンドもまたパケットであるので通常のパケットシーケンス中に多重化することができる。まずCPU10はホストアダプタ16内の特定レジスタに対しREAD/WRITE/COMPARE(リード/ライト/比較)を行いホストアダプタ16が正常であるかどうかを試験する。正常であればホストアダプタ16が使用可能

(5)

特開平5-242006

7

であるためポートアダプタ23に対して“Nop”リクエストパケットを発行するように命令する。“Nop”レスポンスパケットがポートアダプタ23から返ってきたらポートアダプタ23は正常であるので次にディスクアダプタ28に対して“Nop”リクエストパケットを発行するように命令する。“Nop”レスポンスパケットがディスクアダプタ28から返ってきたらディスクアダプタ28は正常である。次に外部記憶デバイスに対して“CHECK”リクエストパケットを発行する。このリクエストは外部記憶デバイスに特別に予約されたCPUのプログラムは使用しない記憶領域に対してREAD/WRITE/COMPAREを行うものである。この結果が正常であればこの一連のチェックに使用したデータバス系は正常であると判断できる。また異常があった場合でもどの部位で異常が起こったかが判断できることになる。なお、図1中のすべての能動エレメントは実際の動作に影響を与えないノーオペレーション(Nop)のようなリクエストに対しては一定時間内にレスポンスを返す機能を持つ。

【0025】以上のように、本実施例によれば、外部記憶装置へのデータ転送を記憶装置の記憶単位の整数倍の長さのデータパケットで行うデータ転送と、パケットはリクエスト系とレスポンス系の2種類があり、指定デバイスへのリクエストに対しては被指定デバイスからのレスポンスの受信により次のリクエストが発行可能となるシーケンスと、装置内のすべての能動エレメントは“ノーオペレーション”リクエストに対しては一定時間内にレスポンスを返す機能とを持つように構成したので、データの転送および制御がパケット形式で行えるためデータ/制御がパケット単位で多重化でき1つのCPUが外部記憶装置を占有しない。またリクエスト系のパケットを発行したCPUはレスポンスのパケットを受信するまでは次のリクエストを発行できないため、外部記憶装置内の制御系はレスポンスの返送を制御することによりCPUの競合を防止できるため明示的なリザーブコマンド/リリースコマンドを廃止することができる。この機能を利用して特定CPUへの優先的な利用を許すこともできる。CPUは外部記憶装置の能動エレメントに対し、“ノーオペレーション”リクエストを発行することによ

8

り、レスポンス受信の結果から該当エレメントおよびそれに係るデータバスのヘルシーチェックを装置の動作に影響なしにオンライン中に実施することができる。

【0026】

【発明の効果】以上のように本発明によれば、当該中央処理装置と外部記憶装置との間のデータ転送を、該記憶装置の記憶単位の整数倍の長さのパケットで行い、該パケットとしてはリクエスト系とレスポンス系を示す2種類のパケットがあり、当該中央処理装置から外部記憶装置の指定記憶デバイスへのリクエストに対しては該指定記憶デバイスからのレスポンスの受信により次のリクエストが当該中央処理装置から発行可能となるように制御するように構成したので、外部記憶装置資源が効率的に利用でき、明示的なリザーブコマンドやリリースコマンドを廃止でき、また、特定の中央処理装置への優先的な外部記憶装置資源の提供が可能になるという効果が得られる。

【図面の簡単な説明】

【図1】この発明の一実施例によるメモリ制御方式を採用した計算機システムの構成を示すブロック図である。

【図2】図1中のホストアダプタの構成を示すブロック図である。

【図3】この実施例において、ホストアダプタからディスクアダプタ間で使用しているパケットの形式を示す図である。

【図4】この実施例においてパケットがどのように実行されるかを示すプロトコルフロー図である。

【図5】この実施例においてノーオペレーションパケットを使ってデータバス系試験を実行している様子を示すプロトコルフロー図である。

【図6】従来のメモリ制御方式を採用した計算機システムの構成を示すブロック図である。

【図7】この従来例において補助記憶装置を制御するCPUのCCWを説明するための図である。

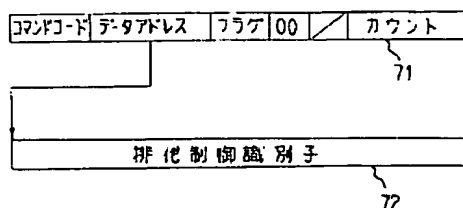
【符号の説明】

10, 11, 12 CPU

30~38 外部記憶デバイス

100 外部記憶装置

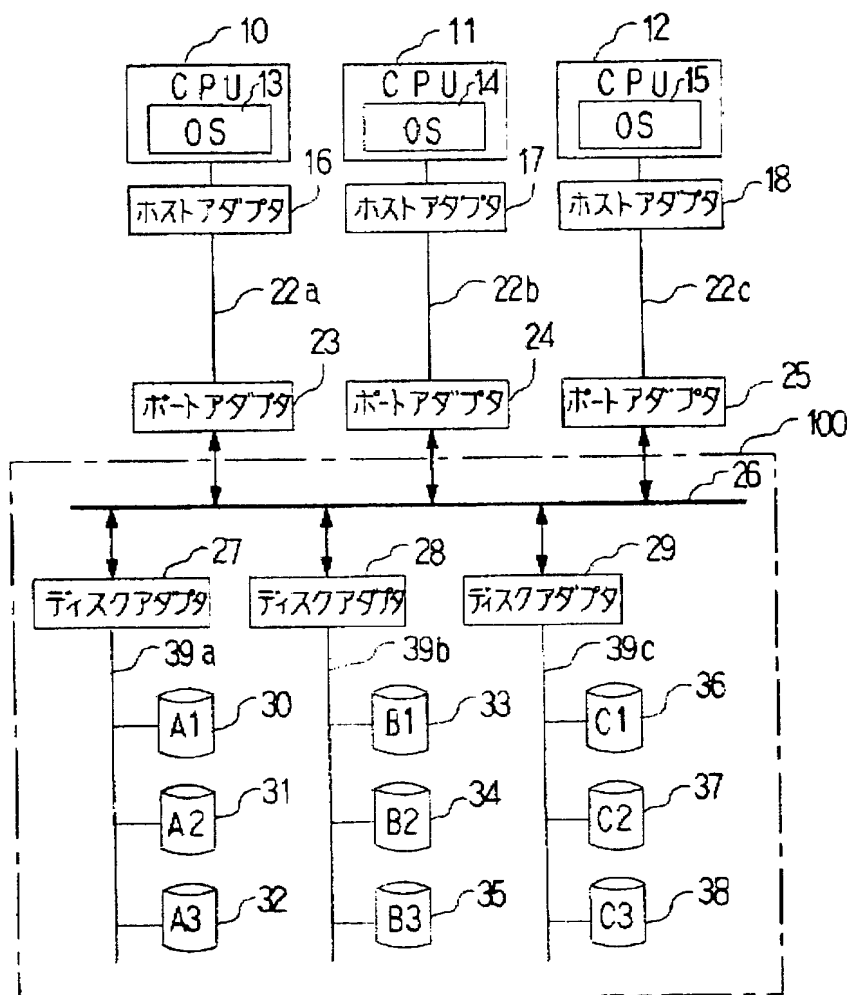
【図7】



(6)

特開平5-242006

【図1】

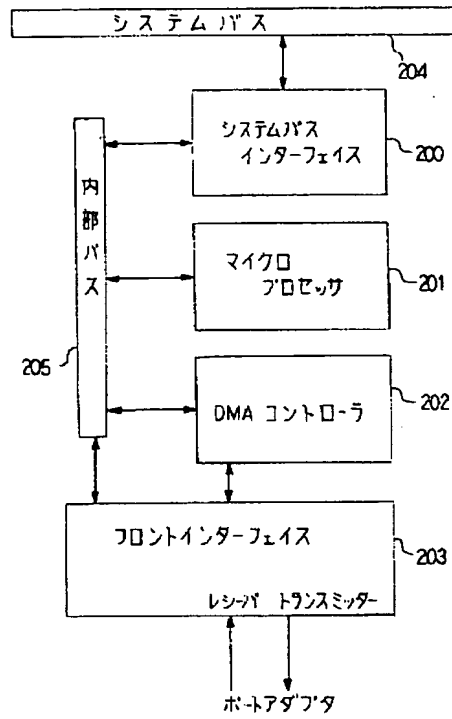


22a~22c; 信号線、26; データバス、
 30~38; 外部記憶デバイス、39a~39c; データバス、
 100; 外部記憶装置

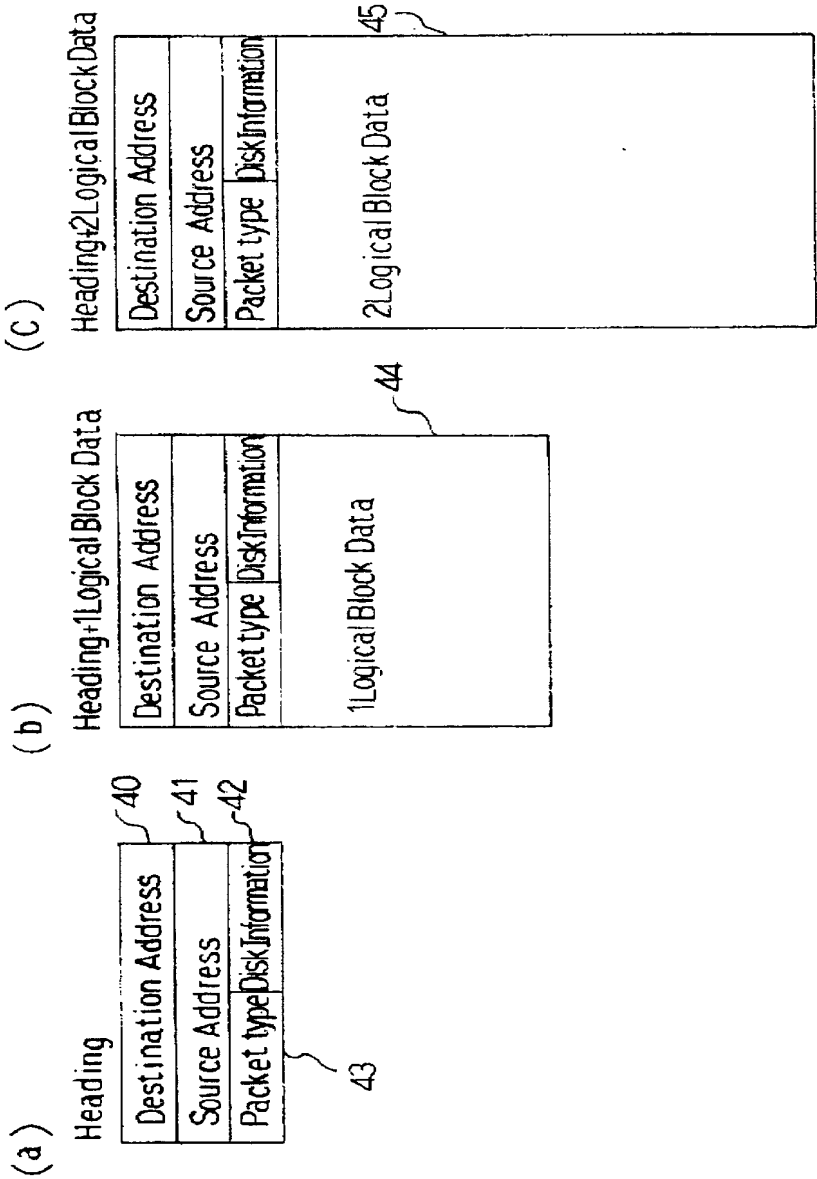
(7)

特開平5-242006

【図2】



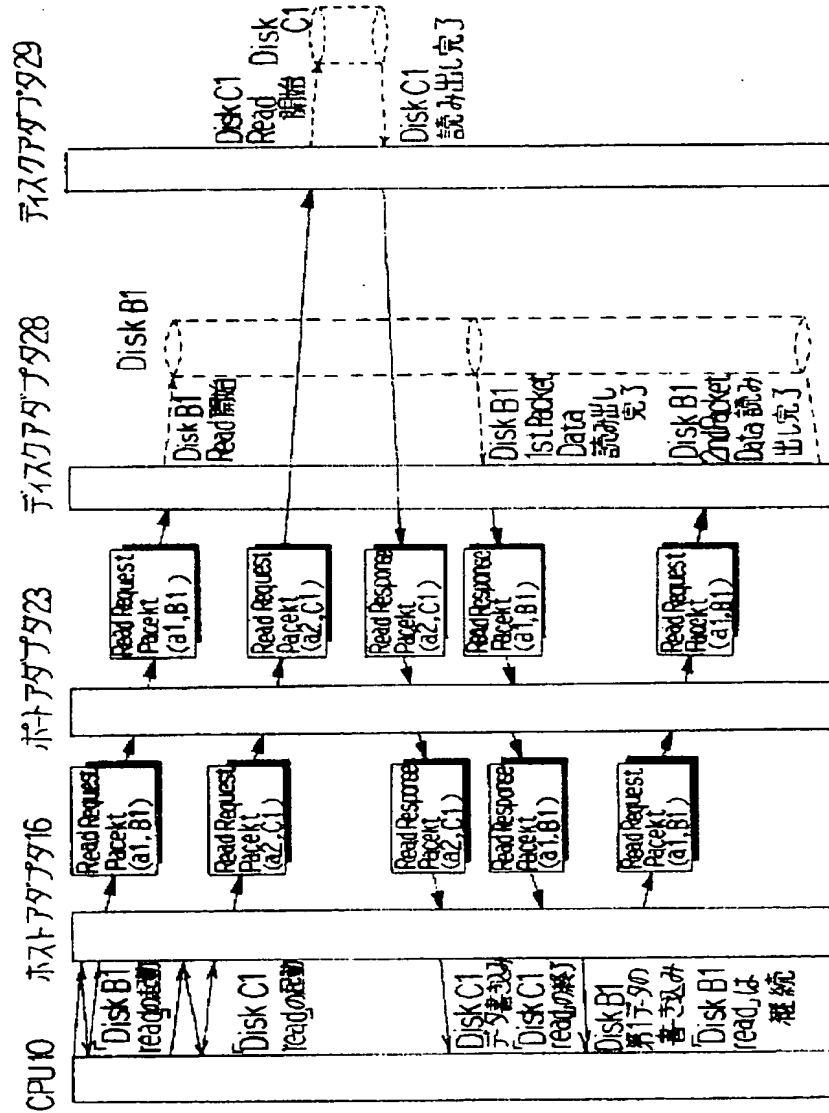
【図3】



(9)

特開平5-242006

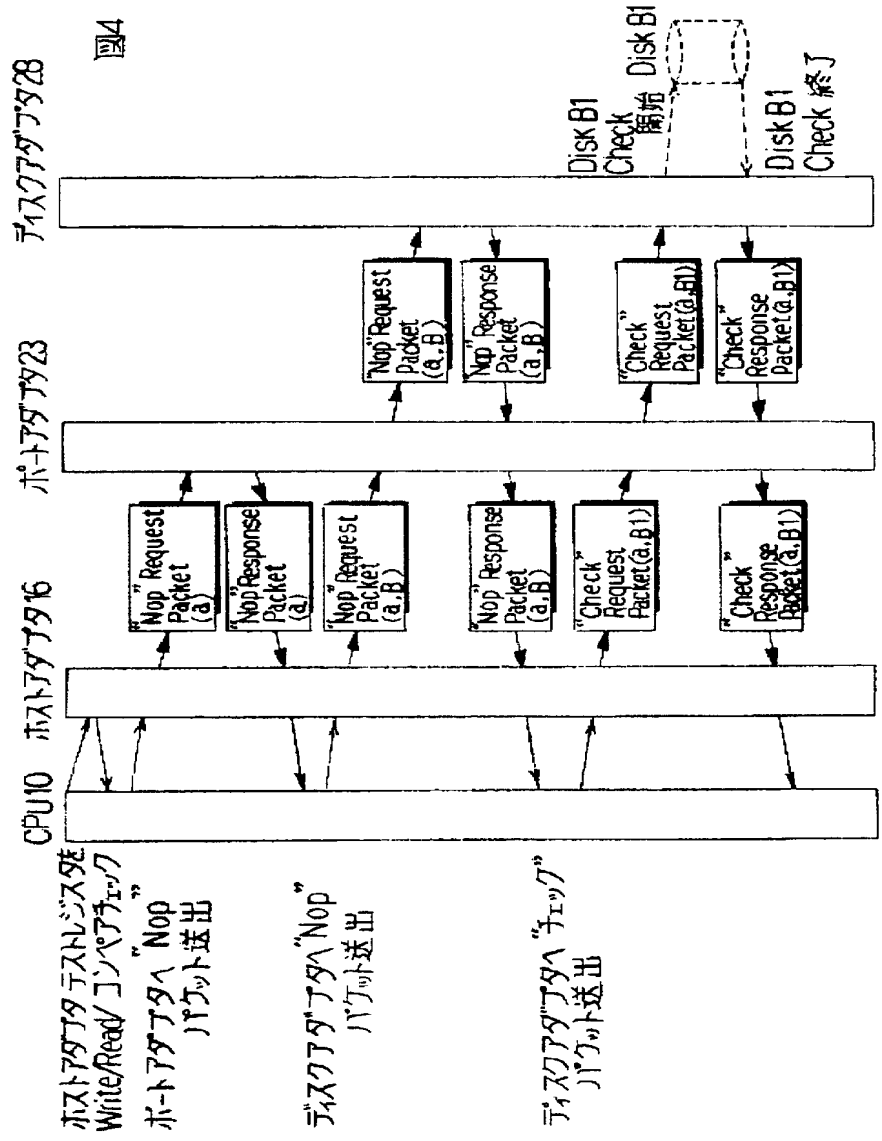
【図 4】



(10)

特開平5-242006

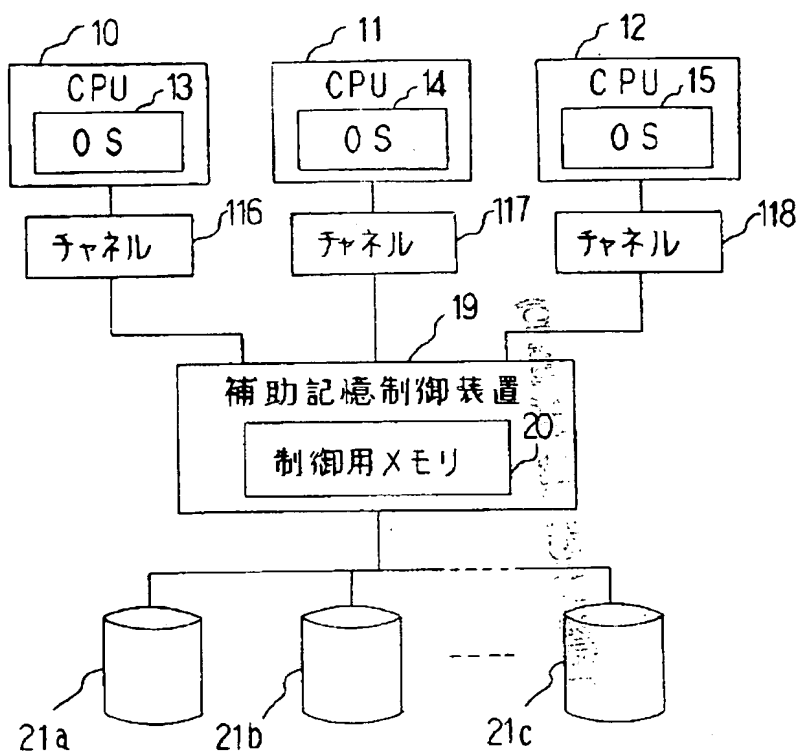
【図5】



(11)

特開平 5-242006

【図 6】



THIS PAGE BLANK (USPTO)